

Binomial upper bounds on generalized moments and tail probabilities of (super)martingales with differences bounded from above

Iosif Pinelis¹

Michigan Technological University

Abstract: Let (S_0, S_1, \dots) be a supermartingale relative to a nondecreasing sequence of σ -algebras $H_{\leq 0}, H_{\leq 1}, \dots$, with $S_0 \leq 0$ almost surely (a.s.) and differences $X_i := S_i - S_{i-1}$. Suppose that $X_i \leq d$ and $\text{Var}(X_i | H_{\leq i-1}) \leq \sigma_i^2$ a.s. for every $i = 1, 2, \dots$, where $d > 0$ and $\sigma_i > 0$ are non-random constants. Let $T_n := Z_1 + \dots + Z_n$, where Z_1, \dots, Z_n are i.i.d. r.v.'s each taking on only two values, one of which is d , and satisfying the conditions $\mathbf{E}Z_i = 0$ and $\text{Var}Z_i = \sigma^2 := \frac{1}{n}(\sigma_1^2 + \dots + \sigma_n^2)$. Then, based on a comparison inequality between generalized moments of S_n and T_n for a rich class of generalized moment functions, the tail comparison inequality

$$\mathbf{P}(S_n \geq y) \leq c \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2}) \quad \forall y \in \mathbb{R}$$

is obtained, where $c := e^2/2 = 3.694\dots$, $h := d + \sigma^2/d$, and the function $y \mapsto \mathbf{P}^{\text{Lin,LC}}(T_n \geq y)$ is the least log-concave majorant of the linear interpolation of the tail function $y \mapsto \mathbf{P}(T_n \geq y)$ over the lattice of all points of the form $nd + kh$ ($k \in \mathbb{Z}$). An explicit formula for $\mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2})$ is given. Another, similar bound is given under somewhat different conditions. It is shown that these bounds improve significantly upon known bounds.

1. Introduction

To begin with, consider normalized Khinchin-Rademacher sums $\varepsilon_1 a_1 + \dots + \varepsilon_n a_n$, where the ε_i 's are i.i.d. Rademacher random variables (r.v.'s), with $\mathbf{P}(\varepsilon_i = \pm 1) = \frac{1}{2}$, and the a_i 's are real numbers such that $a_1^2 + \dots + a_n^2 = 1$. Whittle [27] (cf. Haagerup [10]) established the sharp form

$$(1.1) \quad \mathbf{E} f(\varepsilon_1 a_1 + \dots + \varepsilon_n a_n) \leq \mathbf{E} f\left(\frac{1}{\sqrt{n}}(\varepsilon_1 + \dots + \varepsilon_n)\right) \leq \mathbf{E} f(Z)$$

of Khinchin's inequality [16] for the power moment functions $f(x) = |x|^p$ with $p \geq 3$, where $Z \sim N(0, 1)$. An exponential version of inequality (1.1), with the moment functions $f(x) = e^{\lambda x}$ for $\lambda \in \mathbb{R}$, follows from a result of Hoeffding [12]. An immediate corollary of that is the exponential inequality

$$(1.2) \quad \mathbf{P}(\varepsilon_1 a_1 + \dots + \varepsilon_n a_n \geq x) \leq e^{-x^2/2} \quad \forall x \geq 0.$$

(In fact, Hoeffding [12] obtains more general results; cf. Remark 2.2 below.) This upper bound, $e^{-x^2/2}$, invites a comparison with an "ideal" upper bound, of the

¹Department of Mathematical Sciences, Michigan Technological University, Houghton, Michigan 49931, USA, e-mail: ipinelis@math.mtu.edu

AMS 2000 subject classifications: primary 60E15, 60G42, 60G48, 60G50.

Keywords and phrases: supermartingales, martingales, upper bounds, probability inequalities, generalized moments.

form $cP(Z \geq x)$ for some absolute constant $c > 0$, which one might expect to have in view of (1.1). Since $P(Z \geq x) \sim \frac{1}{x\sqrt{2\pi}}e^{-x^2/2}$ as $x \rightarrow \infty$, this comparison suggests that a factor of order $\frac{1}{x}$ (for large x) is “lost” in (1.2). It turns out that the cause of this loss is that the class of the exponential moment functions is too small. For $\alpha > 0$, consider the following class of functions $f: \mathbb{R} \rightarrow \mathbb{R}$:

$$\mathcal{F}_+^{(\alpha)} := \left\{ f: f(u) = \int_{-\infty}^{\infty} (u-t)_+^{\alpha} \mu(dt) \text{ for some Borel measure } \mu \geq 0 \text{ on } \mathbb{R} \right. \\ \left. \text{and all } u \in \mathbb{R} \right\},$$

where $x_+ := \max(0, x)$ and $x_+^{\alpha} := (x_+)^{\alpha}$. Define $\mathcal{F}_-^{(\alpha)}$ as the class of all functions of the form $u \mapsto f(-u)$, where $f \in \mathcal{F}_+^{(\alpha)}$. Let $\mathcal{F}^{(\alpha)} := \{f + g: f \in \mathcal{F}_+^{(\alpha)}, g \in \mathcal{F}_-^{(\alpha)}\}$.

Proposition 1.1. *For every natural α , one has $f \in \mathcal{F}_+^{(\alpha)}$ iff f has finite derivatives $f^{(0)} := f$, $f^{(1)} := f'$, \dots , $f^{(\alpha-1)}$ on \mathbb{R} such that $f^{(\alpha-1)}$ is convex on \mathbb{R} and $f^{(j)}(-\infty+) = 0$ for $j = 0, 1, \dots, \alpha - 1$.*

The proof of this and other statements (whenever a proof is necessary) is deferred to Section 3.

It follows that, for every $t \in \mathbb{R}$, every $\beta \geq \alpha$, and every $\lambda > 0$, the functions $u \mapsto (u-t)_+^{\beta}$ and $u \mapsto e^{\lambda(u-t)}$ belong to $\mathcal{F}_+^{(\alpha)}$, while the functions $u \mapsto |u-t|^{\beta}$ and $u \mapsto \cosh \lambda(u-t)$ belong to $\mathcal{F}^{(\alpha)}$.

Remark 1.2. Eaton [5] (cf. [7, 21]) obtained inequality (1.1) for a class of moment functions, which essentially coincides with the class $\mathcal{F}^{(3)}$, as seen from [21, Proposition A.1]. Since the class $\mathcal{F}^{(3)}$ is much richer than the class of exponential moment functions, Eaton [6] conjectured (based on asymptotics and numerics) that his inequality (1.1) for $f \in \mathcal{F}^{(3)}$ implies the inequality $P(\varepsilon_1 a_1 + \dots + \varepsilon_n a_n \geq x) \leq \frac{2e^3}{9} \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}$ for all $x > \sqrt{2}$, so that the “lost” factor $\frac{1}{x}$ would be restored. A stronger form of this conjecture was proved by Pinelis [21]:

$$(1.3) \quad P(\varepsilon_1 a_1 + \dots + \varepsilon_n a_n \geq x) \leq \frac{2e^3}{9} P(Z \geq x) \quad \forall x \in \mathbb{R};$$

a multivariate analogue of (1.3) was also obtained there.

Later it was realized (Pinelis [22]) that it is possible to extract (1.3) from (1.1) for all $f \in \mathcal{F}^{(3)}$ because the tail function of the normal distribution is log-concave. The following is a special case of Theorem 4 of Pinelis [23]; see also Theorem 3.11 of Pinelis [22].

Theorem 1.3. *Suppose that $\alpha > 0$, ξ and η are real-valued r.v.’s, and the tail function $u \mapsto P(\eta \geq u)$ is log-concave on \mathbb{R} . Then the comparison inequality*

$$(1.4) \quad Ef(\xi) \leq Ef(\eta) \quad \forall f \in \mathcal{F}_+^{(\alpha)}$$

implies

$$(1.5) \quad P(\xi \geq x) \leq c_{\alpha} P(\eta \geq x) \quad \forall x \in \mathbb{R},$$

where

$$c_{\alpha} := \Gamma(\alpha + 1)(e/\alpha)^{\alpha}.$$

Moreover, the constant factor c_{α} is the best possible one in (1.5).

A similar result for the special case when $\alpha = 1$ is due to Kemperman and is contained in the book by Shorack and Wellner [26, pages 797–799].

Remark 1.4. As follows from [22, Remark 3.13], a useful point is that the requirement of the log-concavity of the tail function $q(u) := \mathbb{P}(\eta \geq u)$ in Theorem 1.3 can be relaxed by replacing q with any (e.g., the least) log-concave majorant of q . However, then the optimality of the constant factor c_α is not guaranteed.

Note that $c_3 = 2e^3/9$, which is the constant factor in (1.3). Bobkov, Götze, and Houdre [4] discovered a simpler proof of a variant of (1.3) with a constant factor $12.0099\dots$ in place of $2e^3/9 = 4.4634\dots$. The value of the constant factor is obviously important in statistical applications. The upper bound in (1.3) improves Chernoff-Hoeffding's bound $e^{-x^2/2}$ in (1.2) for all $x > 1.3124\dots$. On the other hand, the bound in [4] does so only for all $x > 4.5903\dots$, when $\mathbb{P}(Z \geq x) < 2.22 \times 10^{-6}$. The proof in [4] was direct (rather than based on a moment comparison inequality of the form (1.4)). Of course, this does not imply that the direct methods are inferior. In fact, this author has certain ideas to combine the direct and indirect methods to further improve the constant factors.

A stronger, “discrete” version of (1.3) was obtained in [23, Theorem 5], as follows. Let η_1, \dots, η_n be independent zero-mean r.v.'s such that $|\eta_i| \leq 1$ almost surely (a.s.) for all i , and let b_1, \dots, b_n be any real numbers such that $b_1^2 + \dots + b_n^2 = n$. Then

$$(1.6) \quad \mathbb{P}(b_1\eta_1 + \dots + b_n\eta_n \geq x) \leq c_3 \mathbb{P}(\varepsilon_1 + \dots + \varepsilon_n \geq x)$$

for all values x that are taken on by the r.v. $\varepsilon_1 + \dots + \varepsilon_n$ with nonzero probability. Clearly, (1.3) follows from (1.6) by the central limit theorem.

In this paper, we provide new upper bounds on generalized moments and tails of real-valued (super)martingales. It is well known that such bounds can be used, in particular, to obtain concentration-type results; see e.g. [17, 18, 22].

2. Upper bounds on generalized moments and tails of (super)martingales

Theorem 2.1. *Let (S_0, S_1, \dots) be a supermartingale relative to a nondecreasing sequence of σ -algebras $H_{\leq 0}, H_{\leq 1}, \dots$, with $S_0 \leq 0$ a.s. and differences $X_i := S_i - S_{i-1}$, $i = 1, 2, \dots$. Suppose that for every $i = 1, 2, \dots$ there exist non-random constants $d_i > 0$ and $\sigma_i > 0$ such that*

$$(2.1) \quad X_i \leq d_i \quad \text{and}$$

$$(2.2) \quad \text{Var}(X_i | H_{\leq i-1}) \leq \sigma_i^2$$

a.s. Then, for all $n = 1, 2, \dots$,

$$(2.3) \quad \mathbb{E}f(S_n) \leq \mathbb{E}f(T_n) \quad \forall f \in \mathcal{F}_+^{(2)}, \quad \text{where}$$

$$(2.4) \quad T_n := Z_1 + \dots + Z_n$$

and Z_1, \dots, Z_n are independent r.v.'s such that each Z_i takes on only two values, one of which is d_i , and satisfies the conditions

$$\mathbb{E}Z_i = 0 \quad \text{and} \quad \text{Var}Z_i = \sigma_i^2; \quad \text{that is,}$$

$$\mathbb{P}(Z_i = d_i) = \frac{\sigma_i^2}{d_i^2 + \sigma_i^2} \quad \text{and} \quad \mathbb{P}\left(Z_i = -\frac{\sigma_i^2}{d_i}\right) = \frac{d_i^2}{d_i^2 + \sigma_i^2}.$$

Let us explain in general terms how such a result as Theorem 2.1 can be proved. First, it is not difficult to reduce Theorem 2.1 to the case when (S_0, S_1, \dots) is a martingale with $S_0 = 0$ a.s. Then it is not difficult to reduce the situation to the case of one random summand X , so that the problem becomes to find – for any given $t \in \mathbb{R}$, $d > 0$, and $\sigma > 0$ – the maximum of $\mathbb{E}f_t(X)$ subject to the restrictions $X \leq d$ a.s., $\mathbb{E}X = 0$, and $\mathbb{E}X^2 \leq \sigma^2$, where $f_t(x) := (x - t)_+^2$. Then, using arguments going back to Chebyshev and Hoeffding [11, 13, 14, 15], one sees that here an extremal r.v. X takes on at most three distinct values with nonzero probability. In fact, because of a special relation between the objective-function $f_t(x) = (x - t)_+^2$ and the restriction-functions 1 , x , and x^2 , one can see that an extremal r.v. X takes on only two distinct values with nonzero probability; this makes the distribution of an extremal r.v. X uniquely determined by d and σ ; in particular, the extremal distribution does not depend on the value of the parameter t of the objective-function f_t . Thus the result follows.

An alternative approach is based on duality [22, Eq. (4)], and such an approach is actually used in the proof of Theorem 2.1 given in Section 3 below. According to the latter approach, one should search for a tight upper bound on the objective-function $f_t(x)$ over all $x \leq d$; this upper bound must be of the form $Ax^2 + Bx + C$ (which is a linear combination of the restriction-functions 1 , x , and x^2), so that $f_t(x) \leq Ax^2 + Bx + C$ for all $x \leq d$. It is not difficult to see that, for such a tight upper bound, the equality $f_t(x) = Ax^2 + Bx + C$ is attained for at most two distinct values of $x \in (-\infty, d]$. This again implies that an extremal r.v. X takes on only two distinct values with nonzero probability, whence the result. (See the mentioned proof for details.)

Remark 2.2. In the case when (S_i) is a martingale, Theorem 2.1 is a result of Bentkus [1, 3], who used essentially the “duality” approach. Moreover, using Schur convexity arguments similar to those in Eaton [5], he also showed that, in the case $d_i \equiv d$, for every $f \in \mathcal{F}_+^{(2)}$ the right-hand side $\mathbb{E}f(T_n)$ of inequality (2.3) is maximized – under the condition $\sigma_1^2 + \dots + \sigma_n^2 = n\sigma^2 = \text{const}$ – when the Z_i ’s are identically distributed; this fact was earlier established by Hoeffding [12, (2.10)] for $f(x) \equiv e^{\lambda x}$ with $\lambda > 0$. Finally, for such i.i.d. Z_i ’s with $d_i \equiv d$ and $\sigma_i \equiv \sigma$, Bentkus used the method given in [22] (cf. Theorem 1.3 and Remark 1.4 above) to extract the upper bound of the form

$$(2.5) \quad \mathbb{P}(S_n \geq y) \leq c_2 \mathbb{P}^{\text{LC}}(T_n \geq y) \quad \forall y \in \mathbb{R},$$

where the function $y \mapsto \mathbb{P}^{\text{LC}}(T_n \geq y)$ is the least log-concave majorant of the tail function $y \mapsto \mathbb{P}(T_n \geq y)$ on \mathbb{R} . Note that $c_2 = e^2/2 = 3.694\dots$.

Note also that the distribution of r.v. T_n in (2.5) is a shifted and re-scaled binomial distribution, concentrated on the lattice, say $L_{n,d,h}$, of all points of the form $nd + kh$ ($k \in \mathbb{Z}$), where

$$h := d + \sigma^2/d.$$

Here and henceforth, we assume that $d_i \equiv d$, unless indicated otherwise.

Since the tail function of the binomial distribution is log-concave on \mathbb{Z} (see e.g. [23, Remark 13]), one has $\mathbb{P}^{\text{LC}}(T_n \geq y) = \mathbb{P}(T_n \geq y)$ for all y in the lattice $L_{n,d,h}$.

Inequality (2.5) can be significantly improved. Let the function $y \mapsto \mathbb{P}^{\text{Lin}}(T_n \geq y)$ denote the linear interpolation of the function $y \mapsto \mathbb{P}(T_n \geq y)$ over the lattice $L_{n,d,h}$, so that

$$\mathbb{P}^{\text{Lin}}(T_n \geq y) = (1 - \gamma)\mathbb{P}(T_n \geq nd + kh) + \gamma\mathbb{P}(T_n \geq nd + (k + 1)h)$$

if $\gamma := (y - nd - kh)/h \in [0, 1]$ for some $k \in \mathbb{Z}$. Let then the function $y \mapsto \mathbf{P}^{\text{Lin,LC}}(T_n \geq y)$ denote the least log-concave majorant of the function $y \mapsto \mathbf{P}^{\text{Lin}}(T_n \geq y)$ on \mathbb{R} .

Theorem 2.3. *Suppose that the conditions of Theorem 2.1 hold with $d_i \equiv d$. As before, let $\sigma^2 := \frac{1}{n}(\sigma_1^2 + \dots + \sigma_n^2)$ and $h := d + \sigma^2/d$. Then*

$$(2.6) \quad \mathbf{P}(S_n \geq y) \leq c_2 \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \tfrac{h}{2}) \quad \forall y \in \mathbb{R}.$$

Because the tail function $y \mapsto \mathbf{P}(T_n \geq y)$ decreases very rapidly, the shift $\frac{h}{2}$ in $\mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2})$ generally provides quite a substantial improvement. As will be shown later in this paper (Proposition 2.7), in almost all practically important cases the bound (2.6) is better, or even much better, than (2.5).

In Subsection 2.1 (Proposition 2.10), we will also provide an explicit expression for $\mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2})$.

That (S_0, S_1, \dots) in Theorems 2.1 and 2.3 is allowed to be a supermartingale (rather than only a martingale) makes it convenient to use the simple but powerful truncation tool. (Such a tool was used, for example, in [20] to prove limit theorems for large deviation probabilities in Banach spaces based only on precise enough probability inequalities and without using Cramér's transform, the standard device in the theory of large deviations.) Thus, for instance, one immediately has the following corollary of Theorem 2.3.

Corollary 2.4. *Suppose that all conditions of Theorem 2.1 hold except possibly for condition (2.1). Then for all $y \in \mathbb{R}$ and $d > 0$*

$$(2.7) \quad \mathbf{P}(S_n \geq y) \leq \mathbf{P}\left(\max_{1 \leq i \leq n} X_i \geq d\right) + c_2 \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \tfrac{h}{2})$$

$$(2.8) \quad \leq \sum_{1 \leq i \leq n} \mathbf{P}(X_i \geq d) + c_2 \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \tfrac{h}{2}).$$

These bounds are much more precise than the exponential bounds in [8, 9, 19].

Remark 2.5. By the Doob inequality, inequality (2.6) holds for the maximum, $M_n := \max_{0 \leq k \leq n} S_k$, in place of S_n . This follows because (i) all functions of class $\mathcal{F}_+^{(2)}$ are convex and (ii) in view of Lemma 3.1 on page 41, one may assume without loss of generality that (S_i) is a martingale. Similarly, inequalities (2.7) and (2.8) hold for M_n in place of S_n .

In a similar manner, under conditions (2.1)–(2.2) and with

$$(2.9) \quad b := \sqrt{b_1^2 + \dots + b_n^2}, \quad \text{where } b_i := \max(d_i, \sigma_i),$$

Bentkus [2, 3] obtained the following extensions of inequalities (1.6) and (1.3), respectively:

$$(2.10) \quad \mathbf{P}(S_n \geq y) \leq c_3 \mathbf{P}^{\text{LC}}\left(\sum_{i=1}^n \varepsilon_i \geq \frac{y\sqrt{n}}{b}\right) \quad \forall y \in \mathbb{R} \quad \text{and}$$

$$\mathbf{P}(S_n \geq y) \leq c_3 \mathbf{P}(Z \geq y/b) \quad \forall y \in \mathbb{R}.$$

The upper bound in (2.10) can be improved in a similar manner, as follows.

Proposition 2.6. *Under conditions (2.1), (2.2), and (2.9),*

$$(2.11) \quad \mathbf{P}(S_n \geq y) \leq c_3 \mathbf{P}^{\text{Lin,LC}}\left(\sum_{i=1}^n \varepsilon_i \geq 1 + \frac{y\sqrt{n}}{b}\right) \quad \forall y \in \mathbb{R}.$$

In (2.10) and (2.11), the d_i 's are allowed to differ from one another.

Note that the expression $\mathbf{P}^{\text{Lin,LC}}(\sum_{i=1}^n \varepsilon_i \geq 1 + \frac{y\sqrt{n}}{b}) = \mathbf{P}^{\text{Lin,LC}}(\sum_{i=1}^n \frac{b}{\sqrt{n}} \varepsilon_i \geq y + \frac{b}{\sqrt{n}})$ is a special case of the expression $\mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{b}{2})$, with $\frac{b}{2} = d = \sigma = \frac{b}{\sqrt{n}}$.

From the “right-tail” bounds stated above, “left-tail” and “two-tail” ones immediately follow. For instance, if condition $X_i \leq d_i$ in Theorem 2.1 is replaced by $|X_i| \leq d_i$, then inequality (2.3) holds with $\mathcal{F}^{(2)}$ in place of $\mathcal{F}_+^{(2)}$ provided that (S_0, S_1, \dots) is a martingale with $S_0 = 0$ a.s.

In order to present an explicit formula for the upper bound in (2.6) and compare it with the upper bound in (2.5), it is convenient to rescale the r.v. T_n , taking on values in the lattice $L_{n,d,h}$ of all points of the form $nd + kh$ ($k \in \mathbb{Z}$), so that the rescaled r.v., say

$$(2.12) \quad B_n := \frac{q}{d}T_n + np, \quad \text{where} \quad p := \frac{\sigma^2}{d^2 + \sigma^2} \quad \text{and} \quad q := 1 - p = \frac{d^2}{d^2 + \sigma^2},$$

is binomial with parameters n and p . Then for all $y \in \mathbb{R}$, with

$$(2.13) \quad x := \frac{q}{d}y + np,$$

one has

$$(2.14) \quad \begin{aligned} \mathbf{P}(T_n \geq y) &= \mathbf{P}(B_n \geq x) =: Q_n(x), \\ \mathbf{P}^{\text{LC}}(T_n \geq y) &= \mathbf{P}^{\text{LC}}(B_n \geq x) =: Q_n^{\text{LC}}(x), \\ \mathbf{P}^{\text{Lin}}(T_n \geq y) &= \mathbf{P}^{\text{Lin}}(B_n \geq x) =: Q_n^{\text{Lin}}(x), \\ \mathbf{P}^{\text{Lin,LC}}(T_n \geq y) &= \mathbf{P}^{\text{Lin,LC}}(B_n \geq x) =: Q_n^{\text{Lin,LC}}(x), \quad \text{so that} \\ \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2}) &= Q_n^{\text{Lin,LC}}(x + \frac{1}{2}). \end{aligned}$$

Here the function $x \mapsto \mathbf{P}^{\text{Lin,LC}}(B_n \geq x)$ is defined as the least log-concave majorant on \mathbb{R} of the function $x \mapsto \mathbf{P}^{\text{Lin}}(B_n \geq x)$, which is in turn defined as the linear interpolation of the tail function $x \mapsto \mathbf{P}(B_n \geq x)$ over the lattice \mathbb{Z} . Similarly, the function $x \mapsto \mathbf{P}^{\text{LC}}(B_n \geq x)$ is defined as the least log-concave majorant of the function $x \mapsto \mathbf{P}(B_n \geq x)$ on \mathbb{R} .

Note also that $B_n \in [0, n]$ a.s.

Now one is ready to state the following comparison between the upper bounds in (2.5) and (2.6).

Proposition 2.7. *Here relation (2.13) between y and x is assumed.*

(i) *Equation*

$$(2.15) \quad \ln \frac{1-u}{-\ln u} - 1 - \frac{1}{2} \frac{(1+u) \ln u}{1-u} = 0$$

in $u \in (0, 1)$ has exactly one solution, $u = u_ := 0.00505778 \dots$.*

(ii) *The upper bound in (2.6) is no greater than that in (2.5) for all $x \leq j_{**}$, where*

$$(2.16) \quad j_{**} := \left\lfloor \frac{n - u_{**} \frac{q}{p}}{1 + u_{**} \frac{q}{p}} \right\rfloor \quad \text{and}$$

$$(2.17) \quad u_{**} := \frac{u_*}{1 - u_*} = 0.00508349 \dots;$$

*since u_{**} is small, j_{**} is rather close to n unless $\frac{q}{p}$ is large.*

- (iii) Moreover, the upper bound in (2.6) is no greater than that in (2.5) for all $x \leq n$ provided that

$$(2.18) \quad n \leq \frac{p}{q} \frac{1}{u_{**}} = \frac{p}{q} 196.714 \dots$$

In particular, (2.6) works better than (2.5) for all $x \leq n$ if $n \leq 196$ and $p \geq \frac{1}{2}$.

Since $\sum_{i=1}^n \varepsilon_i$ is a special case of T_n (with $d = \sigma = 1$), one immediately obtains the following corollary to Proposition 2.7.

Corollary 2.8. *The upper bound in (2.11) is no greater than that in (2.10) for all $x \leq j_{**}$, where $x := \frac{y\sqrt{n}}{2b} + \frac{n}{2}$ and $j_{**} := \left\lfloor \frac{n-u_{**}}{1+u_{**}} \right\rfloor$. Moreover, the upper bound in (2.11) is no greater than that in (2.10) for all $x \leq n$ provided that $n \leq 196$.*

Of course, the restriction $x \leq j_{**}$ (even though very weak) is only sufficient, but not necessary for the upper bound in (2.6) to be no greater than that in (2.5).

Moreover, (2.6) may work very well even when p is small. For example, in Figure 1 one can see the graph of the ratio

$$r(x) := \frac{Q_n^{\text{Lin,LC}}(x + \frac{1}{2})}{Q_n^{\text{LC}}(x)} = \frac{c_2 \mathbf{P}^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2})}{c_2 \mathbf{P}^{\text{LC}}(T_n \geq y)}$$

of the “new” upper bound – that in (2.6), to the “old” one – that in (2.5), for $n = 30$ and $p = \frac{3}{100}$. In the same figure, one can also see the graph of the new upper bound $q(x) := \min(1, c_2 Q_n^{\text{Lin,LC}}(x + \frac{1}{2}))$, which is a very rapidly decreasing tail function. Proposition 2.7(ii) guaranteed that, for these n and p , the new upper bound will be an improvement of the old one (that is, one will have $r(x) \leq 1$) at least for all $x \in (-\infty, j_{**}] = (-\infty, 25]$, which is rather close to what one can see in the picture. Note that, by Theorem 2.3, for $x \geq 25$ one has $\mathbf{P}(S_n \geq y) \leq c_2 Q_n^{\text{Lin,LC}}(25 + \frac{1}{2}) \approx 3.44 \times 10^{-33}$. Thus, the new upper bound is not an improvement only if the tail probability $\mathbf{P}(S_n \geq y)$ is less than 3.45×10^{-33} . On the other hand, for instance, $r(4) \approx 0.58$; that is, the new upper bound is approximately a 42% improvement of the old one for $x = 4$; at that, the new upper bound is $\approx 0.026 = 2.6\%$, a value quite in a common range in statistical practice.

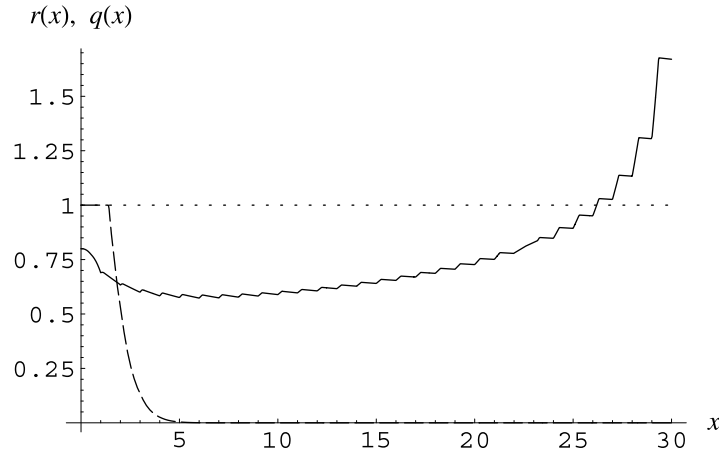


FIG 1. $r(x)$, solid; $q(x) = \min(1, c_2 Q_n^{\text{Lin,LC}}(x + \frac{1}{2}))$, long dashes.

The upper bound in (2.6), $c_2 P^{\text{Lin,LC}}(T_n \geq y + \frac{h}{2}) = c_2 Q_n^{\text{Lin,LC}}(x + \frac{1}{2})$, is rather simple to compute, as described in Proposition 2.10, given in a separate subsection, Subsection 2.1. An underlying reason for this simplicity is that the “discrete” tail function, $\mathbb{Z} \ni j \mapsto Q_n(j)$, of the binomial distribution is log-concave [23, Remark 13]; therefore, it turns out (by Propositions 2.10 and 2.9) that the value $Q_n^{\text{Lin,LC}}(x + \frac{1}{2})$ can be computed locally: for a certain function \mathcal{Q} which depends only on its 7 arguments, one has $Q_n^{\text{Lin,LC}}(x + \frac{1}{2}) = \mathcal{Q}(x, n, j_*, q_{k-1}, q_k, q_{k+1}, q_{k+2})$ for all $x \in \mathbb{R}$ and $p \in (0, 1)$, where $j_* := \lfloor (n+1)p \rfloor + 1$, $k := \lfloor x \rfloor$, and $q_j := Q_n(j)$ for all $j \in \mathbb{Z}$.

2.1. How to compute the expression $Q_n^{\text{Lin,LC}}(x + \frac{1}{2})$ in (2.14)

First here, we need to introduce some notation.

For each $j = 1, \dots, n$ such that $j > (n+1)p$, let

$$(2.19) \quad x_j := j - \frac{1}{2} + \frac{q_j}{p_j} + \frac{\frac{q_j}{p_{j-1}} - \frac{q_j}{p_j}}{\ln \frac{p_{j-1}}{p_j}},$$

$$(2.20) \quad y_j := j - \frac{1}{2} + \frac{q_j}{p_{j-1}} + \frac{\frac{q_j}{p_{j-1}} - \frac{q_j}{p_j}}{\ln \frac{p_{j-1}}{p_j}} = x_j - q_j \left(\frac{1}{p_j} - \frac{1}{p_{j-1}} \right),$$

where, for all $j \in \mathbb{Z}$,

(2.21)

$$q_j := Q_n(j) = P(B_n \geq j) \quad \text{and} \quad p_j := q_j - q_{j+1} = P(B_n = j) = \binom{n}{j} p^j q^{n-j};$$

note that, for each $j = 0, \dots, n$,

$$(2.22) \quad p_{j-1} > p_j \iff j > (n+1)p \iff j \geq j_* := \lfloor (n+1)p \rfloor + 1,$$

so that, for all $j \in \mathbb{Z} \cap [j_*, n]$, one has $\ln \frac{p_{j-1}}{p_j} > 0$, and so, x_j and y_j are well defined, and $x_j > y_j$; note that $j_* \geq 1$. Let also

$$(2.23) \quad x_j := j + \frac{1}{2} \quad \text{and} \quad y_j := j - \frac{1}{2} \quad \text{for integer } j \geq n+1.$$

Proposition 2.9. *For all integer $j \geq j_*$, one has*

$$j - \frac{3}{2} < j - 1 < y_j \leq j - \frac{1}{2} < x_j \leq y_{j+1} \leq j + \frac{1}{2};$$

moreover, if $j \leq n$, then $y_j < j - \frac{1}{2}$ and $x_j < y_{j+1}$.

By Proposition 2.9, for all integer $j \geq j_*$, the intervals

$$\delta_j := (y_j, x_j)$$

are non-empty, with the endpoints

$$y_j \in (j-1, j - \frac{1}{2}] \subset (j - \frac{3}{2}, j - \frac{1}{2}] \quad \text{and} \quad x_j \in (j - \frac{1}{2}, j + \frac{1}{2}];$$

moreover, the intervals δ_j are strictly increasing in $j \geq j_*$: $\delta_j < \delta_{j+1}$ for all $j \geq j_*$, where we use the following convention for any two subsets A and B of \mathbb{R} :

$$A < B \stackrel{\text{def}}{\iff} (x < y \quad \forall x \in A \quad \forall y \in B).$$

For all integer $j \geq j_*$ and all $x \in \delta_j$, introduce the interpolation expression

$$(2.24) \quad Q_n^{\text{Interp}}(x; j) := Q_n^{\text{Lin}}(y_j + \frac{1}{2})^{1-\delta} Q_n^{\text{Lin}}(x_j + \frac{1}{2})^{\delta}, \quad \text{where } \delta := \frac{x - y_j}{x_j - y_j}.$$

Proposition 2.10. *For all real x ,*

$$(2.25) \quad Q_n^{\text{Lin,LC}}\left(x + \frac{1}{2}\right) = Q(x) := \begin{cases} Q_n^{\text{Interp}}(x; j) & \text{if } \exists j \in \mathbb{Z} \cap [j_*, n] \text{ } x \in \delta_j, \\ Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) & \text{otherwise.} \end{cases}$$

A few comments are in order here:

- the function Q in (2.25) is well defined, because the δ_j 's are pairwise disjoint;
- $Q_n^{\text{Lin,LC}}\left(x + \frac{1}{2}\right)$ is easy to compute by (2.25) because, in view of Proposition 2.9, the condition $x \in \delta_j$ for $j \in \mathbb{Z} \cap [j_*, n]$ implies that j equals either $\lfloor x \rfloor$ or $\lfloor x \rfloor + 1$. In particular, $Q_n^{\text{Lin,LC}}\left(x + \frac{1}{2}\right) = Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) = 0$ for all $x \geq n + \frac{1}{2}$.

3. Proofs

Proof of Proposition 1.1. The “only if” part follows because $f \in \mathcal{F}^{(\alpha)}$ for a natural α implies that the (right) derivative $f'(u) = \alpha \int_{-\infty}^{\infty} (u-t)_+^{\alpha-1} \mu(dt)$. Vice versa, if f satisfies the conditions listed after “iff” in the statement of Proposition 1.1, then one can use the Fubini theorem repeatedly to see that for all real u

$$\begin{aligned} f(u) &= \int_{-\infty}^u f'(t) dt = \int_{-\infty}^u dt \int_{-\infty}^t f''(s) ds = \int_{-\infty}^u (u-s) f''(s) ds = \dots \\ &= \int_{-\infty}^u \frac{(u-s)^{\alpha-1}}{(\alpha-1)!} f^{(\alpha)}(s) ds = \int_{-\infty}^u \frac{(u-s)^{\alpha}}{\alpha!} df^{(\alpha)}(s) = \int_{-\infty}^{\infty} (u-s)_+^{\alpha} \mu(ds), \end{aligned}$$

where $f^{(\alpha)}$ is the (nondecreasing) right derivative of the convex function $f^{(\alpha-1)}$ and $\mu(ds) := df^{(\alpha)}(s)/\alpha!$. \square

Theorem 2.1 can be rather easily reduced to the case when (S_n) is a martingale. This is implied by the following lemma.

Lemma 3.1. *Let (S_i) be a supermartingale as in Theorem 2.1, so that conditions (2.1) and (2.2) are satisfied. For $i = 1, 2, \dots$, let*

$$(3.1) \quad \tilde{X}_i := (1 - \gamma_{i-1})X_i + \gamma_{i-1}d_i, \quad \text{where } \gamma_{i-1} := \frac{\mathbb{E}_{i-1}X_i}{\mathbb{E}_{i-1}X_i - d_i};$$

\mathbb{E}_j and Var_j denote, respectively, the conditional expectation and variance given X_1, \dots, X_j . Then \tilde{X}_i is $H_{\leq i}$ -measurable,

$$X_i \leq \tilde{X}_i \leq d_i, \quad \mathbb{E}_{i-1}\tilde{X}_i = 0, \quad \text{and} \quad \text{Var}_{i-1}\tilde{X}_i \leq \text{Var}_{i-1}X_i \leq \sigma_i^2 \quad \text{a.s.}$$

Proof. The conditions $d_i > 0$ and $\mathbb{E}_{i-1}X_i \leq 0$ imply that $\gamma_{i-1} \in [0, 1]$. Now (3.1) and the inequality $X_i \leq d_i$ yield $X_i \leq \tilde{X}_i \leq d_i$. Moreover, (3.1) yields $\mathbb{E}_{i-1}\tilde{X}_i = (1 - \gamma_{i-1})\mathbb{E}_{i-1}X_i + \gamma_{i-1}d_i = 0$ and $\text{Var}_{i-1}\tilde{X}_i = (1 - \gamma_i)^2 \text{Var}_{i-1}X_i \leq \text{Var}_{i-1}X_i$. \square

Theorem 2.1 is mainly based on the following lemma, which also appeared as [1, (12)] and [3, Lemma 4.4(i), with condition $\mathbb{E}X = 0$ missing there].

Lemma 3.2. *Let X be a r.v. such that $\mathbb{E}X = 0$, $\text{Var}X \leq \sigma^2$, and $X \leq d$ a.s. for some non-random constants $\sigma > 0$ and $d > 0$. Let $a := \sigma^2/d^2$, and let X_a be a r.v. taking on values $-a$ and 1 with probabilities $\frac{1}{1+a}$ and $\frac{a}{1+a}$, respectively. Then*

$$(3.2) \quad \mathbb{E}f(X) \leq \mathbb{E}f(d \cdot X_a) \quad \forall f \in \mathcal{F}_+^{(2)}.$$

Proof. (Given here for the reader's convenience and because it is short and simple.) By homogeneity, one may assume that $d = 1$ (otherwise, rewrite the lemma in terms of X/d in place of X). Note that $X_a \leq 1$ with probability 1, $\mathbb{E}X_a = 0$, and $\mathbb{E}X_a^2 = \sigma^2$. Let here $f_t(x) := (x - t)_+^2$ and

$$h_t(x) := \frac{(1 - t)_+^2}{(1 - t_a)^2} (x - t_a)^2, \quad \text{where } t_a := \min(t, -a).$$

Then it is easy to check (by considering the cases $t \geq 1$, $-a \leq t \leq 1$, and $t \leq -a$) that $f_t(x) \leq h_t(x)$ for all $x \leq 1$, and $f_t(x) = h_t(x)$ for $x \in \{1, -a\}$. Therefore, $\mathbb{E}f_t(X) \leq \mathbb{E}h_t(X) \leq \mathbb{E}h_t(X_a) = \mathbb{E}f_t(X_a)$ (the second inequality here follows because $h_t(x)$ is a quadratic polynomial in x with a nonnegative coefficient of x^2 , while $\mathbb{E}X = \mathbb{E}X_a$ and $\mathbb{E}X^2 \leq \mathbb{E}X_a^2$). Now the lemma follows by the definition of the class $\mathcal{F}_+^{(\alpha)}$ and the Fubini theorem. \square

Proof of Theorem 2.1. This proof is based in a standard manner on Lemma 3.2, using also Lemma 3.1. Indeed, by Lemma 3.1, one may assume that $\mathbb{E}_{i-1}X_i = 0$ for all i . Let Z_1, \dots, Z_n be r.v.'s as in the statement of Theorem 2.1, which are also independent of the X_i 's, and let

$$R_i := X_1 + \dots + X_i + Z_{i+1} + \dots + Z_n.$$

Let $\tilde{\mathbb{E}}_i$ denote the conditional expectation given $X_1, \dots, X_{i-1}, Z_{i+1}, \dots, Z_n$. Note that, for all $i = 1, \dots, n$, $\tilde{\mathbb{E}}_i X_i = \mathbb{E}_{i-1} X_i = 0$ and $\tilde{\mathbb{E}}_i X_i^2 = \mathbb{E}_{i-1} X_i^2$; moreover, $R_i - X_i = X_1 + \dots + X_{i-1} + Z_{i+1} + \dots + Z_n$ is a function of $X_1, \dots, X_{i-1}, Z_{i+1}, \dots, Z_n$. Hence, by Lemma 3.2, for any $f \in \mathcal{F}_+^{(2)}$, $\tilde{f}_i(x) := f(R_i - X_i + x)$, and all $i = 1, \dots, n$, one has $\tilde{\mathbb{E}}_i f(R_i) = \tilde{\mathbb{E}}_i \tilde{f}_i(X_i) \leq \tilde{\mathbb{E}}_i \tilde{f}_i(Z_i) = \tilde{\mathbb{E}}_i f(R_{i-1})$, whence $\mathbb{E}f(S_n) \leq \mathbb{E}f(R_n) \leq \mathbb{E}f(R_0) = \mathbb{E}f(T_n)$ (the first inequality here follows because $S_0 \leq 0$ a.s. and any function f in $\mathcal{F}_+^{(2)}$ is nondecreasing). \square

Proof of Theorem 2.3. In view of Theorem 2.1 and Remark 2.2, one has

$$(3.3) \quad \mathbb{E}f(\tilde{S}_n) \leq \mathbb{E}f(B_n) \quad \forall f \in \mathcal{F}_+^{(2)},$$

where $\tilde{S}_n := \frac{q}{d}S_n + np$ and B_n is defined by (2.12). Let U be a r.v., which is independent of B_n and uniformly distributed between $-\frac{1}{2}$ and $\frac{1}{2}$. Then, by Jensen's inequality, $\mathbb{E}f(B_n) \leq \mathbb{E}f(B_n + U)$ for all convex functions f , whence

$$\mathbb{E}f(\tilde{S}_n) \leq \mathbb{E}f(B_n + U) \quad \forall f \in \mathcal{F}_+^{(2)}.$$

Observe that the density function of $B_n + U$ is $x \mapsto \sum_{j=0}^n p_j \mathbf{I}\{|x - j| < \frac{1}{2}\}$ (where the p_j 's are given by (2.21)), and so, the tail function of $B_n + U$ is given by the formula $\mathbb{P}(B_n + U \geq x) = Q_n^{\text{Lin}}(x + \frac{1}{2}) \quad \forall x \in \mathbb{R}$. Now Theorem 2.3 follows by Theorem 1.3, Remark 1.4, and (2.14). \square

Proof of Proposition 2.6. This proof is quite similar to the proof of Theorem 2.3. Instead of (3.3), here one uses inequality $\mathbb{E}f(\tilde{S}_n) \leq \mathbb{E}f(B_n) \quad \forall f \in \mathcal{F}_+^{(3)}$, where $\tilde{S}_n := \frac{\sqrt{n}}{2b}S_n + \frac{n}{2}$. This latter inequality follows from (2.3) (with d_i and σ_i each replaced by $b_i = \max(d_i, \sigma_i)$) and the first one of the inequalities (1.1) $\forall f \in \mathcal{F}_+^{(3)}$ (recall Remark 1.2), taking also into account the inclusion $\mathcal{F}_+^{(3)} \subseteq \mathcal{F}_+^{(2)}$ (which follows e.g. from [23, Proposition 1(ii)]). \square

In the following two propositions, which are immediate corollaries of results of [25] and [24], it is assumed that f and g are differentiable functions on an interval $(a, b) \subseteq (-\infty, \infty)$, and each of the functions g and g' is nonzero and does not change sign on (a, b) ; also, $r := f/g$ and $\rho := f'/g'$.

Proposition 3.1. (Cf. [25, Proposition 1].) *Suppose that $f(a+) = g(a+) = 0$ or $f(b-) = g(b-) = 0$. Suppose also that $\rho \nearrow$ or \searrow ; that is, ρ is increasing or decreasing on (a, b) . Then $r \nearrow$ or \searrow , respectively.*

Proposition 3.2. (Cf. [24, Proposition 1.9]; the condition that $f(a+) = g(a+) = 0$ or $f(b-) = g(b-) = 0$ is not assumed here.) *If $\rho \nearrow$ or \searrow on (a, b) , then $r \nearrow$ or \searrow or $\nearrow\searrow$ or $\searrow\nearrow$ on (a, b) . (Here, for instance, the pattern $\nearrow\searrow$ for ρ on (a, b) means that $\rho \nearrow$ on (a, c) and \searrow on (c, b) for some $c \in (a, b)$; the pattern $\searrow\nearrow$ has a similar meaning.) It follows that, if $\rho \nearrow$ or \searrow or $\nearrow\searrow$ or $\searrow\nearrow$ on (a, b) , then $r \nearrow$ or \searrow or $\nearrow\searrow$ or $\searrow\nearrow$ or $\nearrow\searrow\nearrow$ or $\searrow\nearrow\searrow$ on (a, b) .*

Lemma 3.3. *Part (i) of Proposition 2.7 is true.*

Proof of Lemma 3.3. Let

$$(3.4) \quad h(u) := \ln \frac{1-u}{-\ln u} - 1 - \frac{1}{2} \frac{(1+u) \ln u}{1-u},$$

the left-hand side of (2.15). Here and in rest of the proof of Lemma 3.3, it is assumed that $0 < u < 1$, unless specified otherwise. Then

$$(3.5) \quad h'(u) = r_1(u) := \frac{f_1(u)}{g_1(u)},$$

where $f_1(u) := 2 \ln^2 u + (\frac{1}{u} + 2 - 3u) \ln u + 2(u + \frac{1}{u}) - 4$ and $g_1(u) := -2(1-u)^2 \ln u$. Let next

$$(3.6) \quad r_2(u) := \frac{f'_1(u)}{g'_1(u)} = \frac{f_2(u)}{g_2(u)},$$

where $f_2(u) := (\frac{3}{u} - \frac{1}{u^2}) \ln u + \frac{1}{u} - \frac{1}{u^2}$ and $g_2(u) := 4 \ln u - \frac{2}{u} + 2$, and then

$$(3.7) \quad r_3(u) := \frac{f'_2(u)}{g'_2(u)} = \frac{f_3(u)}{g_3(u)},$$

where $f_3(u) := (2 - 3u) \ln u + 1 + 2u$ and $g_3(u) := 2u(1 + 2u)$.

One has $\frac{f'_3(u)}{g'_3(u)} = -\frac{1}{8}(\frac{2}{u^2} + \frac{3}{u})$, which is increasing; moreover, $\frac{d}{du} \frac{f'_3(u)}{g'_3(u)}$ tends to $-\infty < 0$ and $-29/50 < 0$ as $u \downarrow 0$ and $u \uparrow 1$, respectively. Hence, by Proposition 3.2, $\frac{f'_3(u)}{g'_3(u)}$ is decreasing (in $u \in (0, 1)$).

Next, by (3.7), $r'_3(0+) = \infty > 0$ and $r'_3(1-) = -2/3 < 0$. Hence, by Proposition 3.2, $r_3 \nearrow\searrow$ (on $(0, 1)$).

By (3.6), $r'_2(0+) = \infty > 0$ and $f_2(1) = g_2(1) = 0$. Hence, by Propositions 3.1 and 3.2, $r_2 \nearrow\searrow$ (on $(0, 1)$).

By (3.5), $r'_1(0+) = \infty > 0$ and $r'_1(1-) = -1/4 < 0$. Hence, by Proposition 3.2, $h' = r_1 \nearrow\searrow$ (on $(0, 1)$). Moreover, $h'(0+) = -\infty < 0$ and $h'(1-) = \frac{1}{2} > 0$. Hence, for some $\beta \in (0, 1)$, one has $h' < 0$ on $(0, \beta)$ and $h' > 0$ on $(\beta, 1)$.

Hence, $h \searrow\nearrow$ on $(0, 1)$. Moreover, $h(0+) = \infty$ and $h(1-) = 0$. It follows that the equation $h(u) = 0$ has a unique root $u = u_* \in (0, 1)$.

Now Lemma 3.3 follows by (3.4). \square

Lemma 3.4. *For all $x \leq j_{**} + 1$,*

$$(3.8) \quad Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) \leq Q_n^{\text{LC}}(x).$$

Proof of Lemma 3.4. For any given $x \leq j_{**} + 1$, let

$$j := j_x := \lfloor x \rfloor \quad \text{and} \quad k := k_x := \lfloor x + \frac{1}{2} \rfloor, \quad \text{so that}$$

$$\begin{aligned} j &\leq x < j+1, \quad k - \frac{1}{2} \leq x < k + \frac{1}{2}, \\ Q_n^{\text{LC}}(x) &= q_j^{1-\delta} q_{j+1}^\delta, \quad Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) = (1-\gamma)q_k + \gamma q_{k+1}, \quad \text{where} \\ \delta &:= x - j \in [0, 1) \quad \text{and} \quad \gamma := x + \frac{1}{2} - k \in [0, 1). \end{aligned}$$

There are only three possible cases: $\delta = 0$, $\delta \in [\frac{1}{2}, 1)$, and $\delta \in (0, \frac{1}{2})$.

Case 1: $\delta = 0$. This case is simple. Indeed, here $k = j$ and $\gamma = \frac{1}{2}$, so that

$$Q_n^{\text{LC}}(x) = q_j \geq \frac{1}{2} q_j + \frac{1}{2} q_{j+1} = Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right),$$

since q_j is nonincreasing in j .

Case 2: $\delta \in [\frac{1}{2}, 1)$. This case is simple as well. Indeed, here $k = j + 1$, so that

$$Q_n^{\text{LC}}(x) \geq q_{j+1} \geq (1-\gamma)q_{j+1} + \gamma q_{j+2} = Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right).$$

Case 3: $\delta \in (0, \frac{1}{2})$. In this case, $k = j$ and $\gamma = \delta + \frac{1}{2}$, so that inequality (3.8) can be rewritten here as $q_j^{1-\delta} q_{j+1}^\delta \geq (\frac{1}{2} - \delta)q_j + (\frac{1}{2} + \delta)q_{j+1}$ or, equivalently, as

$$(3.9) \quad F(\delta, u) := u^\delta - \left(\frac{1}{2} - \delta\right) - \left(\frac{1}{2} + \delta\right)u \geq 0,$$

where

$$(3.10) \quad u := \frac{q_{j+1}}{q_j} \in [0, 1];$$

note that the conditions $x \leq j_{**} + 1$ and $j \leq x < j+1$ imply $j \leq j_{**} + 1 \leq n$ (the latter inequality takes place because, by (2.16), $j_{**} < n$); hence, $q_j \geq q_n > 0$, and thus, u is correctly defined by (3.10). Moreover, because both sides of inequality (3.8) are continuous in x for all $x \leq n$ and hence for all $x \leq j_{**} + 1$, it suffices to prove (3.8) only for $x < j_{**} + 1$, whence $j \leq j_{**}$, and so, by (2.16),

$$j \leq \frac{n - u_{**}q/p}{1 + u_{**}q/p};$$

the latter inequality is equivalent, in view of (2.21), to $\frac{p_{j+1}}{p_j} \geq u_{**}$, which in turn implies, in view of (2.17), that

$$\frac{q_j}{q_{j+1}} = 1 + \frac{p_j}{q_{j+1}} \leq 1 + \frac{p_j}{p_{j+1}} \leq 1 + \frac{1}{u_{**}} = \frac{1}{u_*},$$

whence, by (3.10), one obtains $u \geq u_*$. Therefore, the proof in Case 3, and hence the entire proof of Lemma 3.4, is now reduced to the following lemma. \square

Lemma 3.5. *Inequality (3.9) holds for all $\delta \in (0, \frac{1}{2})$ and $u \in [u_*, 1]$.*

Proof of Lemma 3.5. Observe first that, for every $\delta \in (0, \frac{1}{2})$,

$$(3.11) \quad \begin{aligned} F(\delta, 0) &= -\left(\frac{1}{2} - \delta\right) < 0, \quad F(\delta, 1) = 0, \\ (\partial_u F)(\delta, 1) &= -\frac{1}{2} < 0, \text{ and } F \text{ is concave in } u \in [0, 1]. \end{aligned}$$

This implies that, for every $\delta \in (0, \frac{1}{2})$, there exists a unique value

$$u(\delta) \in (0, 1)$$

such that

$$(3.12) \quad F(\delta, u(\delta)) = 0,$$

and at that

$$(3.13) \quad F(\delta, u) \geq 0 \quad \forall \delta \in (0, \frac{1}{2}) \quad \forall u \in [u(\delta), 1],$$

and $(\partial_u F)(\delta, u(\delta))$ is strictly positive and hence nonzero. Thus, equation (3.12) defines an implicit function $(0, \frac{1}{2}) \ni \delta \mapsto u(\delta) \in (0, 1)$. Moreover, since F is differentiable on $(0, \frac{1}{2}) \times (0, 1)$ and $(\partial_u F)(\delta, u(\delta)) \neq 0$ for all $\delta \in (0, \frac{1}{2})$, the implicit function theorem is applicable, so that $u(\delta)$ is differentiable in δ for all $\delta \in (0, \frac{1}{2})$. Now, differentiating both sides of equation (3.12) in δ , one obtains

$$(3.14) \quad u^\delta \ln u + 1 - u + (\delta u^{\delta-1} - \frac{1}{2} - \delta)u'(\delta) = 0,$$

where u stands for $u(\delta)$.

Let us now show that $u(0+) = u(\frac{1}{2}-) = 0$. To that end, observe first that

$$(3.15) \quad \sup_{0 < \delta < \frac{1}{2}} u(\delta) < 1.$$

Indeed, otherwise there would exist a sequence (δ_n) in $(0, \frac{1}{2})$ such that $\varepsilon_n := 1 - u(\delta_n) \downarrow 0$. But then $u(\delta_n)^{\delta_n} = (1 - \varepsilon_n)^{\delta_n} = 1 - \varepsilon_n \delta_n + o(\varepsilon_n)$, so that (3.12) would imply

$$0 = 1 - \varepsilon_n \delta_n + o(\varepsilon_n) - \left(\frac{1}{2} - \delta_n\right) - \left(\frac{1}{2} + \delta_n\right)(1 - \varepsilon_n) = \left(\frac{1}{2} + o(1)\right)\varepsilon_n,$$

which would contradict the fact that $\varepsilon_n = 1 - u(\delta_n) > 0$ for all n .

If it were not true that $u(0+) = 0$, then there would exist a sequence (δ_n) in $(0, \frac{1}{2})$ and some $\varepsilon > 0$ such that $\delta_n \downarrow 0$ while $u(\delta_n) \rightarrow \varepsilon$. But then $u(\delta_n)^{\delta_n} \rightarrow 1$, so that equation (3.12) would imply

$$u(\delta_n) = \frac{u(\delta_n)^{\delta_n} - \left(\frac{1}{2} - \delta_n\right)}{\frac{1}{2} + \delta_n} \rightarrow 1,$$

which would contradict (3.15). Thus, $u(0+) = 0$.

Similarly, if it were not true that $u(\frac{1}{2}-) = 0$, then there would exist a sequence (δ_n) in $(0, \frac{1}{2})$ and some $\varepsilon > 0$ such that $\delta_n \uparrow \frac{1}{2}$ while $u(\delta_n) \rightarrow \varepsilon$. But then equation (3.12) would imply

$$u(\delta_n) = \frac{u(\delta_n)^{\delta_n} - \left(\frac{1}{2} - \delta_n\right)}{\frac{1}{2} + \delta_n} \rightarrow \varepsilon^{1/2},$$

which would imply $0 < \varepsilon = \varepsilon^{1/2}$, so that $\varepsilon = 1$, which would contradict (3.15). Hence, $u(\frac{1}{2}-) = 0$.

Thus, $(0, \frac{1}{2}) \ni \delta \mapsto u(\delta)$ is a strictly positive continuous function, which vanishes at the endpoints 0 and $\frac{1}{2}$. Therefore, there must exist a point $\delta_* \in (0, \frac{1}{2})$ such that $u(\delta_*) \geq u(\delta)$ for all $\delta \in (0, \frac{1}{2})$. Then one must have $u'(\delta_*) = 0$.

Now equation (3.14) yields

$$(3.16) \quad u(\delta_*)^{\delta_*} = -\frac{1 - u(\delta_*)}{\ln u(\delta_*)}, \quad \text{whence}$$

$$(3.17) \quad \delta_* = \frac{\ln \frac{1 - u(\delta_*)}{-\ln u(\delta_*)}}{\ln u(\delta_*)}.$$

In the expression (3.9) for $F(\delta, u)$, replace now u^δ by the right-hand side of (3.16), and then replace δ by the right-hand side of (3.17). Then, recalling (3.12) and slightly re-arranging terms, one sees that $u = u(\delta_*)$ is a root of equation (2.15).

By Lemma 3.3, such a root of (2.15) is unique in $(0, 1)$. It follows that $\max_{0 < \delta < 1/2} u(\delta) = u(\delta_*) = u_* = 0.00505\dots$. In view of (3.13), this completes the proof of Lemma 3.5. \square

Proof of Proposition 2.7. In this proof, we shall use Propositions 2.9 and 2.10 (which will be proved later in this paper) and the following preliminary remarks.

According to [23, Remark 13], the restriction of the tail function Q_n to the set \mathbb{Z} of all integers is log-concave. Therefore, the logarithm, $\ln Q_n^{\text{LC}}$, of the least log-concave majorant Q_n^{LC} of Q_n can be obtained by the linear interpolation of $\ln Q_n$ over \mathbb{Z} , so that

$$(3.18) \quad Q_n^{\text{LC}}(x) = q_j^{1-(x-j)} q_{j+1}^{x-j} \quad \text{if } j \leq x \leq j+1 \text{ \& } j \in \mathbb{Z},$$

where q_j is defined by (2.21). Here and elsewhere, $0^0 := 1$. Recall that the function Q_n^{Lin} is the linear interpolation of the function Q_n over \mathbb{Z} , so that

$$(3.19) \quad Q_n^{\text{Lin}}(x) = (1 - (x - j))q_j + (x - j)q_{j+1} \quad \text{if } j \leq x \leq j+1 \text{ \& } j \in \mathbb{Z}.$$

Since the function Q_n is non-decreasing and left-continuous, one can note that $Q_n^{\text{Lin}} \geq Q_n$ on \mathbb{R} ; also, $Q_n^{\text{Lin}} = Q_n$ on \mathbb{Z} .

Let us now proceed to the proof of Proposition 2.7.

(i) Part (i) of Proposition 2.7 follows by Lemma 3.3.

(ii) If $x \leq j_{**} + 1$ and x is not in δ_j for any $j \in \mathbb{Z} \cap [j_*, n]$, then, by (2.25) and Lemma 3.4,

$$(3.20) \quad Q_n^{\text{Lin,LC}}(x + \tfrac{1}{2}) = Q(x) = Q_n^{\text{Lin}}(x + \tfrac{1}{2}) \leq Q_n^{\text{LC}}(x).$$

If $x \in \delta_j \subset (-\infty, j_{**} + 1]$ for some $j \in \mathbb{Z} \cap [j_*, n]$, then, taking also into account the definition (2.24), (3.20), and the log-concavity of the function Q_n^{LC} , one has, for δ as in (2.24),

$$\begin{aligned} Q_n^{\text{Lin,LC}}(x + \tfrac{1}{2}) &= Q(x) = Q_n^{\text{Interp}}(x; j) \\ &= Q_n^{\text{Lin}}(y_j + \tfrac{1}{2})^{1-\delta} Q_n^{\text{Lin}}(x_j + \tfrac{1}{2})^\delta \leq Q_n^{\text{LC}}(y_j)^{1-\delta} Q_n^{\text{LC}}(x_j)^\delta \leq Q_n^{\text{LC}}(x). \end{aligned}$$

Hence,

$$(3.21) \quad Q_n^{\text{Lin,LC}}(x + \tfrac{1}{2}) \leq Q_n^{\text{LC}}(x)$$

for all $x \leq j_{**} + 1$ except maybe when $x \in \delta_j \cap (-\infty, j_{**} + 1]$ for some $j \in \mathbb{Z} \cap [j_*, n]$ such that $\delta_j \not\subset (-\infty, j_{**} + 1]$. The latter exceptional situation implies that $y_j < j_{**} + 1 < x_j$. Hence, by Proposition 2.9,

$$(3.22) \quad j - 1 < y_j < j_{**} + 1 < x_j \leq j + \frac{1}{2},$$

whence $j < j_{**} + 2$ and $j \geq j_{**} + 1$, so that $j = j_{**} + 1$.

It follows that (3.21) holds (at least) for all

$$x \in (-\infty, j_{**} + 1] \setminus \delta_{j_{**}+1} = (-\infty, y_{j_{**}+1}] \supset (-\infty, j_{**}],$$

the latter inclusion taking place because of the first inequality in (3.22), for $j = j_{**} + 1$. This completes the proof of part (ii) of Proposition 2.7.

(iii) Introduce

$$(3.23) \quad \tilde{Q}(x) := \begin{cases} Q_n^{\text{LC}}(x) & \text{if } x \leq n, \\ q_{n-1}^{n-x} q_n^{x-n+1} & \text{if } x \geq n - 1. \end{cases}$$

Note that for $x \in [n - 1, n]$ the expressions for $\tilde{Q}(x)$ in the two cases in (3.23) coincide with each other, which implies that the function \tilde{Q} is log-concave on \mathbb{R} .

Now we need the following lemma, whose proof will be given a bit later.

Lemma 3.6. *Under condition (2.18), for all $x \in [n, n + \frac{1}{2}]$,*

$$(3.24) \quad Q_n^{\text{Lin}}(x + \tfrac{1}{2}) \leq \tilde{Q}(x).$$

Under condition (2.18), it is easy to see that $j_{**} + 1 \geq n$. Therefore, by Lemma 3.4, one has (3.8) and hence (3.24) for all $x \leq n$.

Also, $Q_n^{\text{Lin}}(x + \frac{1}{2}) = 0$ for all $x \geq n + \frac{1}{2}$.

Thus, by Lemma 3.6, $\tilde{Q}(x) \geq Q_n^{\text{Lin}}(x + \frac{1}{2})$ for all $x \in \mathbb{R}$; at that, as noted above, the function \tilde{Q} is log-concave on \mathbb{R} . Hence, $Q_n^{\text{Lin,LC}}(x + \frac{1}{2}) \leq \tilde{Q}(x)$ for all $x \in \mathbb{R}$. Now part (iii) of Proposition 2.7 follows in view of (3.23), because $\tilde{Q} = Q_n^{\text{LC}}$ on the interval $(-\infty, n]$. \square

Proof of Lemma 3.6. In view of the definitions (3.19) and (3.18), one can rewrite inequality (3.24) as $q_{n-1}^{n-x} q_n^{x-n+1} \geq (n + \frac{1}{2} - x) q_n + (x - n + \frac{1}{2}) q_{n+1}$, for all $x \in [n, n + \frac{1}{2}]$. Since $q_{n+1} = 0$, the latter inequality is equivalent to

$$(3.25) \quad \ln u \leq r(\alpha),$$

where $u := q_{n-1}/q_n > 1$, $\alpha := x - n \in (0, \frac{1}{2})$, and

$$r(\alpha) := \frac{\ln(\frac{1}{2} - \alpha)}{-\alpha}.$$

Since $r'(0+) = -\infty$, $r'(\frac{1}{2}-) = \infty$, and $((\ln(\frac{1}{2} - \alpha))'_\alpha)/(-\alpha'_\alpha) = 1/(\frac{1}{2} - \alpha)$ is increasing in $\alpha \in (0, \frac{1}{2})$, it follows from Proposition 3.2 that there is a unique value $\alpha_* \in (0, \frac{1}{2})$ such that the function r is decreasing on $(0, \alpha_*)$ and increasing on $(\alpha_*, \frac{1}{2})$, so that $r'(\alpha_*) = 0$ and α_* is the point of minimum of function r on $(0, \frac{1}{2})$. In fact, one has $\alpha_* = 0.3133\dots$ and $r(\alpha_*) = 5.3566\dots$.

Therefore, inequality (3.25) can be rewritten as $u \leq e^{r(\alpha_*)}$. On the other hand,

$$u = \frac{q_{n-1}}{q_n} = \frac{np^{n-1}q + p^n}{p^n} = n \frac{q}{p} + 1,$$

so that it suffices to check that $n \frac{q}{p} \leq e^{r(\alpha_*)} - 1 = 211.022\dots$. But this follows from condition (2.18). \square

Proof of Proposition 2.9. Assume the condition $j \geq j_*$. For $j \geq n+1$, the inequalities of Proposition 2.9 immediately follow from (2.23).

It remains to consider the case $j \leq n$, which implies $p_{j-1} > p_j$, by (2.22). Then it suffices to check four inequalities, $j-1 < y_j < j - \frac{1}{2} < x_j < y_{j+1}$. Indeed, the inequality $j - \frac{3}{2} < j-1$ is trivial, and the inequality $y_{j+1} \leq j + \frac{1}{2}$ will then follow — for $j = n$, from (2.23); and for $j = j_*, \dots, n-1$, from the inequalities $y_i < i - \frac{1}{2}$ $\forall i = j_*, \dots, n$.

(i) Checking $j-1 < y_j$. In view of definition (2.20),

$$(3.26) \quad y_j = j - \frac{1}{2} + \kappa_j q_j,$$

where

$$\kappa_j := \frac{1}{p_{j-1}} + \frac{\frac{1}{p_{j-1}} - \frac{1}{p_j}}{\ln \frac{p_{j-1}}{p_j}} = \frac{1 - \frac{p_{j-1}}{p_j} + \ln \frac{p_{j-1}}{p_j}}{p_{j-1} \ln \frac{p_{j-1}}{p_j}},$$

so that

$$(3.27) \quad \kappa_j < 0,$$

in view of the condition $p_{j-1} > p_j$ and the inequality $\ln u < u - 1$ for $u > 1$.

On the other hand, it is well known and easy to verify that the probability mass function (p_k) of the binomial distribution is log-concave, so that the ratio p_k/p_{k-1} is decreasing in k . Hence,

$$q_j = \sum_{k=j}^n p_k \leq \sum_{k=j}^{\infty} p_j \left(\frac{p_j}{p_{j-1}} \right)^{k-j} = \frac{p_j}{1 - \frac{p_j}{p_{j-1}}} =: \hat{q}_j.$$

Therefore, to check $j-1 < y_j$, it suffices to check that $d_j := \hat{y}_j - (j-1) > 0$, where $\hat{y}_j := j - \frac{1}{2} + \kappa_j \hat{q}_j$ (cf. (3.26)). But one can see that $d_j = f(u)/(2(u-1)\ln u)$, where $u := \frac{p_{j-1}}{p_j} > 1$ and $f(u) := 2(1-u) + (1+u)\ln u$. Thus, to check $j-1 < y_j$, it suffices to show that $f(u) > 0$ for $u > 1$. But this follows because $f(1) = f'(1) = 0$ and f is strictly convex on $(1, \infty)$.

(ii) Checking $y_j < j - \frac{1}{2}$. This follows immediately from (3.26) and (3.27).

(iii) Checking $j - \frac{1}{2} < x_j$. This follows because $x_j - (j - \frac{1}{2}) = q_j(\ln u - 1 + 1/u)/(p_j \ln u) > 0$, where again $u := \frac{p_{j-1}}{p_j} > 1$.

(iv) Checking $x_j < y_{j+1}$. Let first $j \leq n-1$, so that $p_j > p_{j+1} > 0$. In view of (2.20) and the obvious identity $1 + \frac{q_{j+1}}{p_j} = \frac{q_j}{p_j}$, one has

$$y_{j+1} = j - \frac{1}{2} + \frac{q_j}{p_j} + \frac{\frac{q_{j+1}}{p_j} - \frac{q_{j+1}}{p_{j+1}}}{\ln \frac{p_j}{p_{j+1}}},$$

so that the inequality $x_j < y_{j+1}$, which is being checked, can be rewritten as $q_j r_j > q_{j+1} r_{j+1}$ or, equivalently, as

$$(3.28) \quad \sum_{k=0}^{\infty} (p_{j+k} r_j - p_{j+1+k} r_{j+1}) > 0,$$

where $r_j := (\frac{1}{p_j} - \frac{1}{p_{j-1}})/\ln \frac{p_{j-1}}{p_j}$. Note that $r_j > 0$, $r_{j+1} > 0$, and $p_j r_j = h(v) := (1-v)/(-\ln v)$, where $v := p_j/p_{j-1} \in (0, 1)$. By Proposition 3.1, $h(v)$ is increasing in $v \in (0, 1)$. On the other hand, $v = p_j/p_{j-1}$ is decreasing in j , by the mentioned

log-concavity of (p_j) . It follows that $p_j r_j$ is decreasing in j . Because of this and the same log-concavity, $\frac{p_{j+k} r_j}{p_{j+1+k} r_{j+1}} > \frac{p_j r_j}{p_{j+1} r_{j+1}} > 1 \quad \forall k = 0, 1, \dots$, which yields (3.28).

Finally, in the case when $j = n \geq j_*$, the inequality $x_j < y_{j+1}$ follows from (2.19) and (2.23), because then

$$x_n = n + \frac{1}{2} + \frac{\frac{p_n}{p_{n-1}} - 1}{\ln \frac{p_{n-1}}{p_n}} < n + \frac{1}{2} = y_{n+1}.$$

□

Proof of Proposition 2.10.

Step 1. Here we observe that the function Q defined in (2.25) is continuous. Indeed, the function $x \mapsto Q_n^{\text{Lin}}(x + \frac{1}{2})$ is defined and continuous everywhere on \mathbb{R} . On the other hand, for every integer $j \geq j_*$, the function $x \mapsto Q_n^{\text{Interp}}(x; j)$ is defined and continuous on the interval δ_j ; moreover, it continuously interpolates on the interval δ_j between the values of the function $x \mapsto Q_n^{\text{Lin}}(x + \frac{1}{2})$ at the endpoints, y_j and x_j , of the interval δ_j . Also, the intervals δ_j with $j \in \mathbb{Z} \cap [j_*, n]$ are pairwise disjoint. Thus, the function Q is continuous everywhere on \mathbb{R} .

Step 2. Here we show that the function Q is log-concave. To that end, introduce

$$\ell_j(x) := \ln\left(\left(\frac{1}{2} + j - x\right)q_j + \left(\frac{1}{2} - j + x\right)q_{j+1}\right) \quad \forall j \in \mathbb{Z}, \quad \text{so that}$$

$$(3.29) \quad \left(j - \frac{1}{2} \leq x < j + \frac{1}{2} \text{ \& } j \leq n\right) \implies \ln Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) = \ell_j(x);$$

here, the condition $j \leq n$ provides for both sides of the equality in (3.29) to be defined. One can check (which is better done using Mathematica or similar software) the basic relations

$$(3.30) \quad j \in \mathbb{Z} \cap [j_*, n] \implies \ell'_{j-1}(y_j) = \ell'_j(x_j) = \frac{\ell_j(x_j) - \ell_{j-1}(y_j)}{x_j - y_j};$$

these relations do not rely on the fact that the q_j 's pertain to a binomial distribution, but only on general relations: $p_{j-1} > p_j > 0$, $p_i = q_i - q_{i+1} \quad \forall i$, and $q_{j+1} \geq 0$, as well as the inequalities $(j-1) - \frac{1}{2} < y_j < (j-1) + \frac{1}{2}$ and $j - \frac{1}{2} < x_j \leq j + \frac{1}{2}$, which follow by Proposition 2.9 and ensure that ℓ_j and ℓ_{j-1} are defined and differentiable in neighborhoods of x_j and y_j , respectively. Using the latter relations together with (2.24) and (3.29), one has

$$(3.31) \quad \frac{d}{dx} \ln Q_n^{\text{Interp}}(x; j) = \frac{\ell_j(x_j) - \ell_{j-1}(y_j)}{x_j - y_j} \quad \forall x \in \delta_j \quad \forall j \in \mathbb{Z} \cap [j_*, n].$$

Moreover, for all integer $j \leq n$,

$$\ell'_j\left(j - \frac{1}{2}\right) = \frac{q_{j+1} - q_j}{q_j} = \frac{-p_j}{q_j} \quad \text{and} \quad \ell'_{j-1}\left(j - \frac{1}{2}\right) = \frac{-p_{j-1}}{q_j}.$$

Hence and by (2.22), for every integer $j \leq n$ one has

$$(3.32) \quad \ell'_j\left(j - \frac{1}{2}\right) \leq \ell'_{j-1}\left(j - \frac{1}{2}\right) \iff j \leq j_* - 1.$$

In view of (3.29), the function $x \mapsto \ln Q_n^{\text{Lin}}(x + \frac{1}{2})$ is concave on the interval $[j - \frac{1}{2}, j + \frac{1}{2}]$ for every integer $j \leq n$. Note also that $\bigcup_{j \leq j_* - 1} [j - \frac{1}{2}, j + \frac{1}{2}] = (-\infty, j_* - \frac{1}{2}]$. Hence, by (3.32) and (3.29), the function

$$(3.33) \quad x \mapsto \ln Q_n^{\text{Lin}}\left(x + \frac{1}{2}\right) \quad \text{is concave on the interval } (-\infty, j_* - \frac{1}{2}].$$

In addition to the open intervals $\delta_j = (y_j, x_j)$, introduce the closed intervals

$$\Delta_j := [x_j, y_{j+1}]$$

for integer $j \geq j_*$. Then, by Proposition 2.9 and (2.23), the intervals Δ_j are each nonempty,

$$(3.34) \quad \delta_{j_*} \cup \Delta_{j_*} \cup \delta_{j_*+1} \cup \Delta_{j_*+1} \cup \cdots \cup \delta_n \cup \Delta_n = (y_{j_*}, n + \tfrac{1}{2}],$$

and

$$\delta_{j_*} < \Delta_{j_*} < \delta_{j_*+1} < \Delta_{j_*+1} < \cdots < \delta_n < \Delta_n.$$

Thus, the intervals δ_j and Δ_j with $j \in \mathbb{Z} \cap [j_*, n]$ form a partition of the interval $(y_{j_*}, n + \frac{1}{2}]$. Moreover, for every $j \in \mathbb{Z} \cap [j_*, n]$, by Proposition 2.9, $\Delta_j \subseteq [j - \frac{1}{2}, j + \frac{1}{2}]$, and so, by (3.29), the function

$$(3.35) \quad x \mapsto \ln Q_n^{\text{Lin}}(x + \tfrac{1}{2}) \quad \text{is concave on the interval } \Delta_j.$$

Also, by (2.24), for every $j \in \mathbb{Z} \cap [j_*, n]$, the function

$$(3.36) \quad x \mapsto \ln Q_n^{\text{Interp}}(x; j) \quad \text{is concave (in fact, affine) on the interval } \delta_j.$$

By the definition of Q in (2.25), for all $x \in \mathbb{R}$ and all $j \in \mathbb{Z} \cap [j_*, n]$, one has

$$(3.37) \quad Q(x) = \begin{cases} Q_n^{\text{Lin}}(x + \tfrac{1}{2}) & \text{if } x \leq y_{j_*}, \\ Q_n^{\text{Interp}}(x; j) & \text{if } x \in \delta_j \text{ \& } j \in \mathbb{Z} \cap [j_*, n], \\ Q_n^{\text{Lin}}(x + \tfrac{1}{2}) & \text{if } x \in \Delta_j \text{ \& } j \in \mathbb{Z} \cap [j_*, n], \\ 0 = Q_n^{\text{Lin}}(x + \tfrac{1}{2}) & \text{if } x \geq n + \tfrac{1}{2}. \end{cases}$$

Note also that, by Proposition 2.9, $y_{j_*} \leq j_* - \frac{1}{2}$, so that $(-\infty, y_{j_*}] \subseteq (-\infty, j_* - \frac{1}{2}]$. Now it follows from (3.37), (3.33), (3.35), and (3.36) that the function $\ln Q$ is concave on each of the disjoint adjacent intervals

$$(3.38) \quad (-\infty, y_{j_*}], \delta_{j_*}, \Delta_{j_*}, \delta_{j_*+1}, \Delta_{j_*+1}, \dots, \delta_n, \Delta_n,$$

whose union is the interval $(-\infty, n + \frac{1}{2}]$. Moreover, it follows from the continuity of Q (established in Step 1) and formulas (3.30), (3.29), and (3.31) that the function $\ln Q$ is differentiable at all the endpoints $y_{j_*}, x_{j_*}, y_{j_*+1}, x_{j_*+1}, \dots, y_n, x_n$ of the intervals (3.38) except the right endpoint $y_{n+1} = n + \frac{1}{2}$ of the interval Δ_n .

Therefore, the function $\ln Q$ is concave on the interval $(-\infty, n + \frac{1}{2})$. On the other hand, $\ln Q = -\infty$ on the interval $[n + \frac{1}{2}, \infty)$. Thus, it is proved that the function $\ln Q$ is concave everywhere on \mathbb{R} .

Step 3. Here we show that

$$(3.39) \quad Q(x) \geq Q_n^{\text{Lin}}(x + \tfrac{1}{2})$$

for all real x . In view of (3.37) and (3.34), it suffices to check (3.39) for $x \in \delta_j$ with $j \in \mathbb{Z} \cap [j_*, n]$. By Proposition 2.9, $\delta_j \subseteq (j - 1, j + \frac{1}{2}] \subset [j - \frac{3}{2}, j + \frac{1}{2}]$, for every $j \in \mathbb{Z} \cap [j_*, n]$.

By (3.29), the function $x \mapsto \ln Q_n^{\text{Lin}}(x + \frac{1}{2}) = \ell_j(x)$ is concave on the interval $[j - \frac{1}{2}, j + \frac{1}{2}]$, for every integer $j \leq n$. Hence, (3.30) and (2.24) imply that, for all $x \in \delta_j \cap [j - \frac{1}{2}, j + \frac{1}{2}]$ with $j \in \mathbb{Z} \cap [j_*, n]$,

$$\begin{aligned} \ln Q_n^{\text{Lin}}(x + \tfrac{1}{2}) &= \ell_j(x) \leq \ell_j(x_j) + \ell'_j(x_j)(x - x_j) \\ &= \frac{x_j - x}{x_j - y_j} \ell_{j-1}(y_j) + \frac{x - y_j}{x_j - y_j} \ell_j(x_j) = \ln Q_n^{\text{Interp}}(x; j) = \ln Q(x), \end{aligned}$$

so that one has (3.39) for all $x \in \delta_j \cap [j - \frac{1}{2}, j + \frac{1}{2}]$ with $j \in \mathbb{Z} \cap [j_*, n]$. Similarly (using inequality $\ell_{j-1}(x) \leq \ell_{j-1}(y_j) + \ell'_{j-1}(y_j)(x - y_j)$) it can be shown that (3.39) takes place for all $x \in \delta_j \cap [j - \frac{3}{2}, j - \frac{1}{2}]$ with $j \in \mathbb{Z} \cap [j_*, n]$. This completes Step 3.

Step 4. Here we show that, if \tilde{Q} is a log-concave function on \mathbb{R} such that

$$(3.40) \quad \tilde{Q}(x) \geq Q_n^{\text{Lin}}(x + \tfrac{1}{2}) \quad \forall x \in \mathbb{R},$$

then $\tilde{Q} \geq Q$ on \mathbb{R} . In view of (3.37), it suffices to check that $\tilde{Q} \geq Q$ on δ_j for every $j \in \mathbb{Z} \cap [j_*, n]$. But, by (3.40), one has $\tilde{Q}(y_j) \geq Q_n^{\text{Lin}}(y_j + \frac{1}{2})$ and $\tilde{Q}(x_j) \geq Q_n^{\text{Lin}}(x_j + \frac{1}{2})$. Hence, taking into account the log-concavity of \tilde{Q} and (2.24) and (3.37), one has, for all $x \in \delta_j$ with $j \in \mathbb{Z} \cap [j_*, n]$ and δ as in (2.24),

$$\tilde{Q}(x) \geq \tilde{Q}(y_j)^{1-\delta} \tilde{Q}(x_j)^\delta \geq Q_n^{\text{Lin}}(y_j + \tfrac{1}{2})^{1-\delta} Q_n^{\text{Lin}}(x_j + \tfrac{1}{2})^\delta = Q_n^{\text{Interp}}(x; j) = Q(x).$$

The facts established in Steps 2, 3, and 4 imply that the function Q is indeed the least log-concave majorant of the function $x \mapsto Q_n^{\text{Lin}}(x + \frac{1}{2})$. Thus, Proposition 2.10 is proved. \square

References

- [1] BENTKUS, V. (2002). A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Lithuanian Math. J.* **42** 262–269. MR1947624
- [2] BENTKUS, V. (2003). An inequality for tail probabilities of martingales with differences bounded from one side. *J. Theoret. Probab.* **16** 161–173. MR1956826
- [3] BENTKUS, V. (2004). On Hoeffding's inequalities. *Ann. Probab.* **32** 1650–1673. MR2060313
- [4] BOBKOV, S. G., GÖTZE, F. AND HOUDRÉ, C. (2001). On Gaussian and Bernoulli covariance representations. *Bernoulli* **7** 439–451. MR1836739
- [5] EATON, M. L. (1970). A note on symmetric Bernoulli random variables. *Ann. Math. Statist.* **41** 1223–1226. MR268930
- [6] EATON, M. L. (1974). A probability inequality for linear combinations of bounded random variables. *Ann. Statist.* **2** 609–614.
- [7] FIGIEL, T., HITCZENKO, P., JOHNSON, W. B., SCHECHTMAN, G. AND ZINN, J. (1997). Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities. *Trans. Amer. Math. Soc.* **349** 997–1027. MR1390980
- [8] FUK, D. H. (1971). Certain probabilistic inequalities for martingales. *Siberian Math. J.* **14** 131–137. MR0293695
- [9] FUK, D. H. AND NAGAEV, S. V. (1971). Probabilistic inequalities for sums of independent random variables (in Russian, English summary). *Teor. Veroyatnost. i Primenen.* **16** 660–675. MR0293695
- [10] HAAGERUP, U. (1982). The best constants in the Khinchine inequality. *Studia Math.* **70** 231–283. MR0654838
- [11] Hoeffding, W. (1955). The extrema of the expected value of a function of independent random variables. *Ann. Math. Statist.* **26** 268–275. MR70087
- [12] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR144363
- [13] Hoeffding, W. AND SHRIKHANDE, S. S. (1955). Bounds for the distribution function of a sum of independent, identically distributed random variables. *Ann. Math. Statist.* **26** 439–449. MR72377

- [14] KARLIN, S. AND STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Pure and Applied Mathematics, Vol. XV. Interscience Publishers John Wiley & Sons, New York–London–Sydney. MR204922
- [15] KARR, A. F. (1983). Extreme points of certain sets of probability measures, with applications. *Math. Oper. Res.* **8** (1) 74–85. MR703827
- [16] KHINCHIN, A. (1923). Über dyadische Brüche. *Math. Z.* **18** 109–116.
- [17] MCDIARMID, C. (1989). On the method of bounded differences. In *Surveys in Combinatorics, 1989 (Norwich, 1989)*. London Math. Soc. Lecture Note Ser., Vol. **141**. Cambridge Univ. Press, Cambridge, pp. 148–188. MR1036755
- [18] MCDIARMID, C. (1998). Concentration. In *Probabilistic methods for algorithmic discrete mathematics*. Algorithms Combin., Vol. **16**. Springer, Berlin, pp. 195–248. MR1678578
- [19] NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7** 745–789. MR0542129
- [20] PINELIS, I. F. (1985). Asymptotic equivalence of the probabilities of large deviations for sums and maximum of independent random variables (in Russian). *Limit Theorems of Probability Theory. Trudy Inst. Mat.* **5**. Nauka Sibirsk. Otdel., Novosibirsk, pp. 144–173, 176. MR0821760
- [21] PINELIS, I. (1994). Extremal probabilistic problems and Hotelling’s T^2 test under a symmetry condition. *Ann. Statist.* **22** (1), 357–368. MR1272088
- [22] PINELIS, I. (1998). Optimal tail comparison based on comparison of moments. *High Dimensional Probability (Oberwolfach, 1996)*. *Progr. Probab.* **43**. Birkhäuser, Basel, pp. 297–314. MR1652335
- [23] PINELIS, I. (1999). Fractional sums and integrals of r -concave tails and applications to comparison probability inequalities. *Advances in Stochastic Inequalities (Atlanta, GA, 1997)*, *Contemp. Math.* **234**. Amer. Math. Soc., Providence, RI, pp. 149–168. MR1694770
- [24] PINELIS, I. (2001). L’Hospital type rules for oscillation, with applications. *JIPAM. J. Inequal. Pure Appl. Math.* **2** (3) Article 33, 24 pp. (electronic). MR1876266
- [25] PINELIS, I. (2002). L’Hospital type results for monotonicity, with applications. *JIPAM. J. Inequal. Pure Appl. Math.* **3** (1) Article 5, 5 pp. (electronic). MR1888920
- [26] SHORACK, G. R. AND WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York. MR0838963
- [27] WHITTLE, P. (1960). Bounds for the moments of linear and quadratic forms in independent variables. *Teor. Veroyatnost. i Primenen.* **5** 331–335. MR0133849